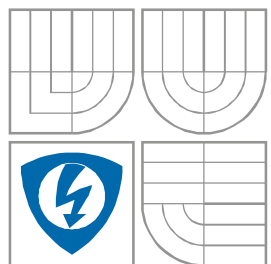


**VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ**  
BRNO UNIVERSITY OF TECHNOLOGY



**FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH  
TECHNOLOGIÍ**  
**ÚSTAV RADIOELEKTRONIKY**

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION  
DEPARTMENT OF RADIO ELECTRONICS

# **TVORBA A ZPRACOVÁNÍ SIGNÁLOVÉ DATABÁZE**

CREATION AND PROCESSING OF SIGNAL DATABASE

**DIPLOMOVÁ PRÁCE**  
MASTER'S THESIS

**AUTOR PRÁCE**  
AUTHOR

Bc. Jiří Glett

**VEDOUCÍ PRÁCE**  
SUPERVISOR

prof. Ing. Milan Sigmund, CSc.

BRNO, 2009



**VYSOKÉ UČENÍ  
TECHNICKÉ V BRNĚ**

**Fakulta elektrotechniky  
a komunikačních technologií**

**Ústav radioelektroniky**

# Diplomová práce

magisterský navazující studijní obor  
**Elektronika a sdělovací technika**

**Student:** Bc. Jiří Glett

**ID:** 86631

**Ročník:** 2

**Akademický rok:** 2008/2009

## NÁZEV TÉMATU:

### **Tvorba a zpracování signálové databáze**

## POKYNY PRO VYPRACOVÁNÍ:

Seznamte se s metodami vytváření signálových databází, jejich strukturování a používání. Vytvořte přehledovou studii o dostupných databázích hudebních a řečových signálů. Pojednejte o speciálních softwarových prostředcích na tvorbu databází. Navrhnete strukturu pro vlastní databázi.

Vypracujte českou verzi anglické databáze SUSAS a obdobně vytvořte novou databázi, přičemž část signálových záznamů vám bude dodána a část pořídíte samostatně.

Vytvořenou databázi doplňte o ukázkové metody zpracování řečových signálů včetně získaných vzorových výsledků. Celkovou databázi realizujte v české a anglické verzi.

## DOPORUČENÁ LITERATURA:

[1] PSUTKA, J., MÜLLER, Z., MATOUŠEK, J., RADOVÁ, V. Mluvíme s počítačem česky. Praha: Academia, 2006.

[2] GIBBON, D., MOORE, R., WINSKI, R. Handbook of Standards and Resources for Spoken Language Systems, New York: Mouton de Gruyter, 1997.

**Termín zadání:** 9.2.2009

**Termín odevzdání:** 29.5.2009

**Vedoucí práce:** prof. Ing. Milan Sigmund, CSc.

**prof. Dr. Ing. Zbyněk Raida**

*Předseda oborové rady*

## UPOZORNĚNÍ:

Autor diplomové práce nesmí při vytváření diplomové práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení § 152 trestního zákona č. 140/1961 Sb.

## ABSKTRAKT

Práce stručně seznamuje s historií a vznikem databází celkově. Rozebírá metodiku strukturování, třídění a účel použití. Práce se dále zabývá konkrétními databázovými softwary určenými přímo ke zpracování audio signálů. Dále pojednává o programech, které umožňují vytváření vlastní databázové struktury. Práce také pojednává o databázi SUSAS a rozebírá její obsah. Je vytvořena vlastní databáze hudebních signálů, která obsahuje několik skupin hudebních záznamů podobných vždy v určitém hledisku. Řečová databáze obsahuje řečové záznamy z databáze SUSAS a dále nahrávky z televizních pořadů, reality-show, sportovních přenosů, reportáží a dokumentů, kdy mluvčí je vystaven stresu a emocím. Výsledkem práce je databázový program, který všechny záznamy účinně třídí a zpracovává. Databázi je možné volně rozšiřovat. Výsledný program byl zrealizován v české i anglické verzi.

## ABSTRACT

The work become acquainted with history and rise databases globally. Construes philosophy structuralization, sorting and purpose using. Work further deal with concrete databases softwares intended directly to processing audio signals. Further treat of programmes that the make possible generation personal database structures. Work deal about SUSAS database and analysing its content. It is created self database of music signals, which includes several musical groups always similar in certain aspect. Speech database contains records from the SUSAS database and records from television programs, reality-shows, sports broadcasts, reports and documents, when speakers are subjected to stress and emotions. The work is a database program that can effectively classify all records and processes. The database can be freely extended. The resulting program was realized into Czech and English version.

## KLÍČOVÁ SLOVA

databáze, struktura, databázový software, zvuk, zpracování, stres, melodie

## KEYWORDS

database, structure, database software, audio, processing, stress, melody

GLETT, J. Tvorba a zpracování signálové databáze. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2009. 38 s., 1.příl. Vedoucí diplomové práce prof. Ing. Milan Sigmund, CSc.

## **Prohlášení**

Prohlašuji, že svou diplomovou práci na téma Tvorba a zpracování signálové databáze jsem vypracoval samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené diplomové práce dále prohlašuji, že v souvislosti s vytvořením této diplomové práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení § 152 trestního zákona č. 140/1961 Sb.

V Brně dne 29. května 2009

.....  
podpis autora

## **Poděkování**

Děkuji vedoucímu diplomové práce Prof. Ing. Milanu Sigmundovi, CSc. za účinnou metodickou, pedagogickou a odbornou pomoc a další cenné rady při zpracování mé diplomové práce.

V Brně dne 29. května 2009

.....  
podpis autora

## Obsah:

1. Úvod .....	8
2. Historie .....	8
3. Databázové modely.....	9
4. Databázová integrita .....	9
5. Multimediální databáze.....	9
6. Audio databázové programy .....	10
6.1 Oracle Database.....	10
6.2 Weddle's Audio Info DataBase .....	10
6.3 Music Database 2000 .....	11
6.4 AudioList.....	11
6.5 Catraxe .....	11
7. Software k vytváření vlastních databází: .....	12
7.1 Brilliant Database.....	12
7.2 PHP Magic .....	13
7.3 Web Database Server .....	13
10. Program Signálová Databáze .....	14
10.1 Úvod.....	14
10.2 Hudební databáze .....	14
10.3 Melodie různými nástroji .....	14
10.4 Stejná píseň ve více žánrech.....	16
10.5 Řečová databáze.....	17
10.6 Databáze řeči pod stresem.....	17
10.7 Databáze českých mluvčích .....	19
10.8 SUSAS .....	20
10.9 Spektrogramy a časové průběhy.....	20
10.10 Jazykové rozhraní.....	21
10.11 Instalace databázového serveru .....	21
10.12 Vytvoření databáze.....	21

10.13 Přidávání zvukových záznamů do databáze .....	22
10.14 Informace .....	23
11. Nahrávání a generování zvuků .....	23
12. Formát souborů v databázi .....	23
13. Všeobecné audio formáty .....	24
13.1 Lidské vnímání zvuku .....	24
13.2 Vlastní audio formáty .....	25
14. Česká dokumentace pro databázi SUSAS .....	34
15. Převod souborů databáze SUSAS .....	34
15.1 Sph convert.....	34
15.2 Sound Converter.....	35
16. Závěr .....	36
17. Seznam literatury .....	37
18. Zdroje zvukových signálů a použité hudební programy.....	38

## 1. Úvod

Databáze (neboli Datová základna) je určitá uspořádaná množina informací (dat) uložená na paměťovém médiu. V širším smyslu jsou součástí databáze i softwarové prostředky, které umožňují manipulaci s uloženými daty a přístup k nim. Tento systém se v české odborné literatuře nazývá systém řízení báze dat. Běžně se označením databáze – v závislosti na kontextu – myslí jak uložená data, tak i software. [2]

## 2. Historie

Předchůdcem databází byly papírové kartotéky. Umožňovaly uspořádávání dat podle různých kritérií a zařizování nových položek. Veškeré operace s nimi prováděl přímo člověk. Správa takových kartoték byla v mnohém podobná správě dnešních databází.

Dalším krokem bylo převedení zpracování dat na stroje. Za první velké strojové zpracování dat lze asi považovat sčítání lidu ve Spojených státech v roce 1890. Paměťovým médiem byl děrný štítek a zpracování sebraných informací probíhalo na elektromechanických strojích. Elektromechanické stroje se využívaly pro účely zpracování dat další půlstoletí.

Velkým impulsem pro další rozvoj databází byl překotný vývoj počítačů v padesátých letech 20. století. Ukázalo se, že původně univerzální používání strojového kódu procesorů je (nejen) pro databázové úlohy neefektivní, a proto se objevil požadavek na vyšší jazyk pro zpracování dat.

V roce 1959 se konala konference zástupců firem, uživatelů a amerického ministerstva obrany, jejímž závěrem byl požadavek na univerzální databázový jazyk. Výsledkem byla o rok později na konferenci CODASYL publikovaná první verze jazyka COBOL, který byl po mnoho dalších let nejrozšířenějším jazykem pro hromadné zpracování dat.

V roce 1965 na konferenci CODASYL byl vytvořen výbor Database Task Group, který měl za úkol vytvořit koncepci databázových systémů. Začaly vznikat první síťové SŘBD na sálových počítačích. Jedním z prvních průkopníků databází byl Charles Bachman.

V roce 1971 vydal výbor zprávu The DBTG April 1971 Report, kde se objevily pojmy jako schéma databáze, jazyk pro definici schématu, subschéma a podobně. Byla zde popsána celá architektura síťového databázového systému.

Ve stejné době byly vyvíjeny i hierarchické databáze. Jedním z prvních SŘBD byl IMS, který byl vyvinut firmou IBM pro program letu na Měsíc Program Apollo. Systém IMS patří stále k nejrozšířenějším na sálových počítačích.

V roce 1970 začínají zveřejnění článků E. F. Codd první relační databáze, které pohlíží na data jako na tabulky. Kolem roku 1974 se vyvíjí první verze dotazovacího jazyka SQL. Vývoj této technologie po 10 letech přinesl výkonově použitelné systémy, srovnatelné se síťovými a hierarchickými databázemi.

V 90. letech 20. století se začínaly objevovat první objektově orientované databáze, jejichž filozofie byla přebírána z objektově orientovaných jazyků. Tyto databáze měly podle předpokladů vytlačit relační systémy. Původní předpoklady se však nenaplnily a vznikla kompromisní objektově-relační technologie. [4]



### **3. Databázové modely**

Z hlediska způsobu ukládání dat a vazeb mezi nimi můžeme rozdělit databáze do základních typů:

- Hierarchická databáze
- Síťová databáze
- Relační databáze
- Objektová databáze
- Objektově relační databáze [2]

### **4. Databázová integrita**

Databázová integrita je takový stav, při němž záznamy v celé databázi vyhovují soustavě určitých definovaných pravidel. Tato pravidla obvykle odpovídají vybraným pravidlům z té části světa, pro kterou databáze slouží. [2]

### **5. Multimediální databáze**

Multimediální databáze (resp. databázový systém) je databázový systém, který dokáže spravovat multimediální data, jako jsou v našem případě audio soubory.

Multimediální data jsou nestrukturovaná data vyznačující se velkým objemem (jedna položka může mít desítky až tisíce megabajtů). Příkladem jsou 2D obrázky, 3D modely, video, audio, dokumenty a jejich kombinace. Tato data obsahují často také metadata (např. exif u obrázků, ID3 u mp3).

Dotazování a získávání informací z kolekcí uložených dat může probíhat podle popisu nebo obsahu.

Příkladem databázového dotazu v takovémto systému je nalezení nejpodobnějších obrázků k danému obrázku (tzv. podobnostní vyhledávání). [2]

## 6. Audio databázové programy

### 6.1 Oracle Database

<http://www.oracle.com>

Oracle Multimedia (dříve Oracle *interMedia*) se věnuje databázím a vydává je ve standardní a podnikatelské edici.

- To je hlavní účelový rys, který umožňuje účinné zpracování a získávání obrazových, zvukových, a video dat.
- Oracle Multimedia má znalost nejpopulárnějších multimediálních formátů a může automaticky vyjmout metadata a umožnit základní zpracování.
- Oracle Multimedia zvyšuje produktivitu programátora při rozvíjení multimediálních aplikací používáním JSPs, Servlets, PL/SQL nebo při používání nástrojů jako JDeveloper (ADF/UIX) a Oracle Portal.

Oracle Multimedia řídí multimediální data uvnitř Oracle databáze pod transakční kontrolou. Střídavě, Oracle databáze může ukládat a indexovat meta- informace společně s externími adresami, které umožní přístup k obsahu uloženému vně databáze.

Cena:

standardní verze:     £ 152.0

podnikatelská verze: £ 405.0

### 6.2 Weddle's Audio Info DataBase

<http://weddlesoftware.com/AudioInfoDB.htm>

Weddle's Audio Info DataBase je jednoduchý program, díky kterému lze ukládat informace o jakékoliv hudební sbírce. Uchovává informace o umělcích, albech, písních, textech, délkách souborů, producentech, poznámkách a dalších. Weddle's Audio Info DataBase nepodporuje automatickou četbu CD nebo stahování mimo internetového.

Cena: 8.15 €

### **6.3 Music Database 2000**

<http://www.musicdatabasesoftware.com>

Software k organizaci vaší hudební sbírky nebo hudební knihovny. Pro DJ's, hudebníky, sběratele, rozhlasové stanice, estrádní umělci, hudebníky, záznamové obchody.

Výhodná věc je že užívá CDDb (TM) Databáze. Tudíž sám vyhledá celé tituly, takže se nemusí psát ručně.

CDDb (TM) Databázové statistiky: Přes 2.5 milionů CD, přes 30 milionů písní, 30 milionů uživatelských přístupů za měsíc. A novinka je, že můžete také vyhledávat alba bez CD. Užívá Microsoft(R) přístup '97 - standardní databázový formát. Výrobci vědí, že nemá smysl vytvářet vlastní formáty jako další databázové programy. Má také kategorii pro BPM (počet úderů za minutu) užívaný hlavně DJi, bohužel CDDb nemá zatím mnoho BPM záznamů.

Cena: \$ 39.95

### **6.4 AudioList**

<http://www.wakefieldsoft.com/audiolist/>

AudioList je kompletní softwarový organizér audio inventáře. Kontroluje vaši sbírku hudby, kterou aktuálně vlastníte, co jste slyšeli, co máte půjčené, co si chcete koupit, atd. Program je dostupný ve 4 verzích:

- Palma OS® a Windows PC
- Pocket PC nebo Windows mobile PC
- Windows PC jen
- Palma OS® jen, určený pro Mac

Cena: \$29.99

### **6.5 Catraxe**

<http://www.filesland.com/companies/FNProgramware/CATraxe.html>

Catraxe je schopný databázový program navržený k tomu, aby vám pomohl zkatalogizovat a řídit vaši hudební sbírku. Je možné zkatalogizovat veškeré mediální druhy, včetně CD, LP, kazet, minidisků a mp3. Podporuje CDDb - světově největší hudební internetovou databázi. Je určený pro Windows prostředí. Uživatelské rozhraní je komplexní, intuitivní a snadno použitelné. Komplexní soubor dat umožňuje zkatalogizovat široký okruh informací: umělec, titul, mediální formát, rok vydání, štítek, číslo katalogu, kupní cena, aktuální hodnota, hudební kategorie, hrací čas, názvy tracků, skladatelé písní, producenti, nahrávací studia, hudebníci atd. Lze přiložit booklet, připojit mp3 soubor ke každému tracku i vaše osobní komentáře. Můžete efektivně najít informace, které hledáte použitím filtrování. Program pravidelně kontroluje kdo z vašich přátel co má půjčeno. Vždy budete vědět, kdo má co a od které doby. Jednotka statistik vám dovolí analyzovat vaši sbírku různými způsoby. Catraxe vám eviduje vaši počáteční investici a celkovou hodnotu vaší sbírky.

Cena: \$ 39.0

## 7. Software k vytváření vlastních databází:

### 7.1 Brilliant Database

<http://www.brilliantdatabase.com/>

Briliant je databázový softwarový program, který vám dovolí vytvořit databázi bez toho, mít nějakou zvláštní znalost nebo schopnosti programování. Použitím tohoto programu můžete snadno navrhovat vaši vlastní databázi:

- **Formy.** Je zde 22 různých polí pracujících s texty, daty, vzorci, obrazy.
- **Styly.** Navrhuje zprávy s grafikou, grafy, kalkulacemi uživatelskými poli a tabulkami.
- **Dotazy.** S vizuálním dotazovacím editorem, netřeba učit se dotazovací jazyk.
- **Matematické vzorce.** Automatizuje veškerou matematickou kalkulaci v databázových polích a zprávách.
- **Skripty.** Použitím vizuálního skript-editoru můžete automatizovat vaši práci s databází bez studování zvláštního programovacího jazyka.
- **Uživatelské rozhraní.** Přidejte nové tlačítka, menu, „horké“ klávesy, grafiku, vytvořte uživatelské akce... .
- **Uživatelský export/import.** Vytváří šablony k tomu, aby exportovaly data do Webu, textového dokumentu, schránky a propojte vaši databázi s MS Word a E-Mail software.
- **Navrhuje váš databázový software.** Zvláštní SDK Edice dovoluje sestavit databázi do nezávislé exe - aplikace, a potom její distribuce je volná, bez příplatků a na neomezené množství počítačů.

Každá databáze vytvořená za pomoci Brilliant Database podporuje síť, stromovou a vzdálenou strukturu, registruje přílohy, 4Gb databáze, import/export atd.

Cena:

komerční verze: \$149

vzdělávací verze: \$79

## 7.2 PHP Magic

<http://www.websitedatabases.com/>

Databázový software pro rozvoj internetových databází. Lze vytvořit webovou databázi bez dřívějších znalostí PHP nebo MySQL kódování.

Vytvoříte databázové tabulky a pole a program už sám vygeneruje celý kód. Jednoduše nahrajte vygenerovaný kód na váš server a spusťte vygenerovaný nastavovací soubor ve vašem prohlížeči. A je to! Budete mít vytvořenou profesionální databázovou aplikaci včetně dat navigace, třídění, hledání, pokročilého filtrování, zaznamenávání dodatku, modifikace, mazání, souboru a další rysy. Možnosti jak použít webovou databázi jsou nekonečné. PHPmagic to dělá všechno pro vás, automaticky.

Cena: \$49,95

## 7.3 Web Database Server

<http://www.lightdb.com>

PrimaSoft Web Database Server je internetová aplikace která vám umožní rychle vytvořit, řídit a podělit se o vaši databázi skrz vaši společnost nebo s přáteli přes Internet nebo Intranet. Můžete nastavit konkrétní databázi pro váš obchod, databázi znalostí pro vaši organizaci, nebo katalog produktů pro vaši firemní webovou stránku. Můžete vytvořit databázi informací pro vaše osobní použití.

Web Database Server je dostupný ve dvou variantách:

### **Web dB Server:**

Web database server je základní verzí programu.

### **dB Organizer Deluxe:**

Database Organizer Deluxe je rozšířená verze programu, která kromě základní verze umožňuje vaši databázi exportovat vašemu webovému databázovému serveru. Ten už je dobře připravený na online zpracování.

Cena: \$65

## 10. Program Signálová Databáze

### 10.1 Úvod

Program Signálová databáze je určen jako výukový program pro názorné ukázky hudebních a řečových signálů. Program se spouští souborem signalDb.exe.

Jeho hlavní okno se skládá ze tří částí:

- Stromová struktura: nachází se vlevo nahoře a zobrazuje kompletní stromovou strukturu aktuální databáze. Pomocí kurzorových kláves a případně kláves + a – procházíme databází do nižších úrovní až ke zvoleným ukázkám.
- Pole s ukázkami zvolené kategorie: nachází se pod stromovou strukturou a zobrazuje všechny audio ukázky zvolené sekce. Všechny hudební ukázky v databázi je možno přehrát pomocí tlačítka se symbolem play. Zastavit pomocí tlačítka stop. Pokud se hudební ukázka přehrává, není možné jakkoliv jinak pracovat s programem. Program dále umožňuje uložit jakoukoliv skladbu pomocí tlačítka Ulož (uloží se i adekvátní spektrogram a časový průběh). Zatrhávací políčko slouží k zobrazení spektrogramu a časového průběhu dané ukázky.
- Největší část programu zabírá plocha pro zobrazení spektrogramů a časových průběhů. Současně je možné zobrazit tato grafická znázornění až ke třem audio ukázkám.

### 10.2 Hudební databáze

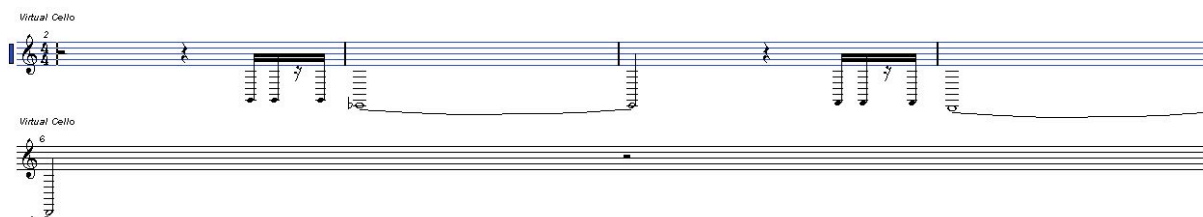
V hudební části se nachází ukázky Melodie různými nástroji, kde jsou tři známé melodie hrány na různé nástroje v různých tóninách, reálnými i virtuální nástroji, či různými technikami hraní. V těchto ukázkách lze sledovat určité souvislosti. Dále jsou zde ukázky stejné písně ve více žánrech. Tyto ukázky prezentují stejnou píseň, avšak zahrnou jiným umělcem, nebo je zde prezentace rozdílu studiové a „live“ - koncertní verze. V těchto ukázkách můžeme sledovat rovněž určitou souvislost.

### 10.3 Melodie různými nástroji

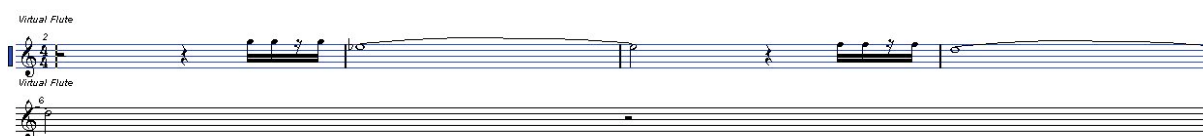
Jako ukázkovou melodii jsem zvolil úvodních několik tónů Beethovenovy 5. symfonie, 1. věty. Nahrál jsem ji následujícími reálnými nástroji: Basa, flétna, kytara, violoncello, viola a housle.

Stejnou melodii jsem zaznamenal notovým záznamem a nechal přehrát virtuálními midi nástroji. Vybíral jsem adekvátní virtuální nástroje, aby byly zvukově podobné předešlým reálným nástrojům.

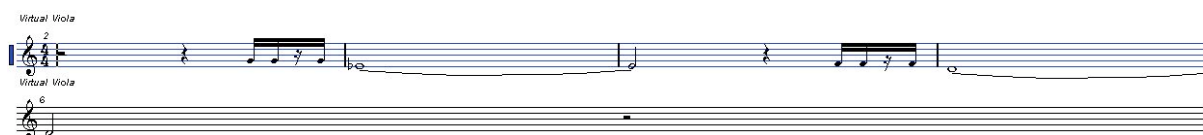
Demonstrativně zde mám několik notových partitur k virtuálním nástrojům. Notový záznam pro živý nástroj je samozřejmě stejný, rozdíl je v tom, že virtuální nástroj je strojově přesný, zatímco živý nástroj je zkrátka „živý“, není syntetický a v jeho záznamu je zanesena stopa muzikanta. Kvalitní hudební nástroj bude vždy lepší než sebelepší virtuální nástroj.



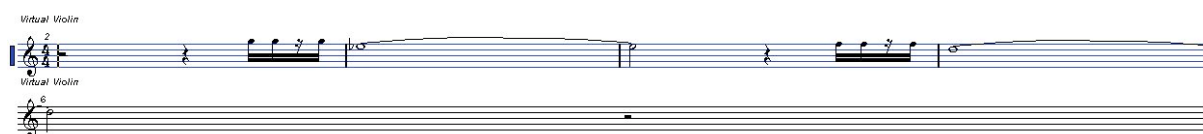
**Obr.1.: Notová partitura pro virtuální nástroj – violoncello.**



**Obr.2.: Notová partitura pro virtuální nástroj – flétna.**



**Obr.3.: Notová partitura pro virtuální nástroj – viola.**



**Obr.4.: Notová partitura pro virtuální nástroj – housle.**

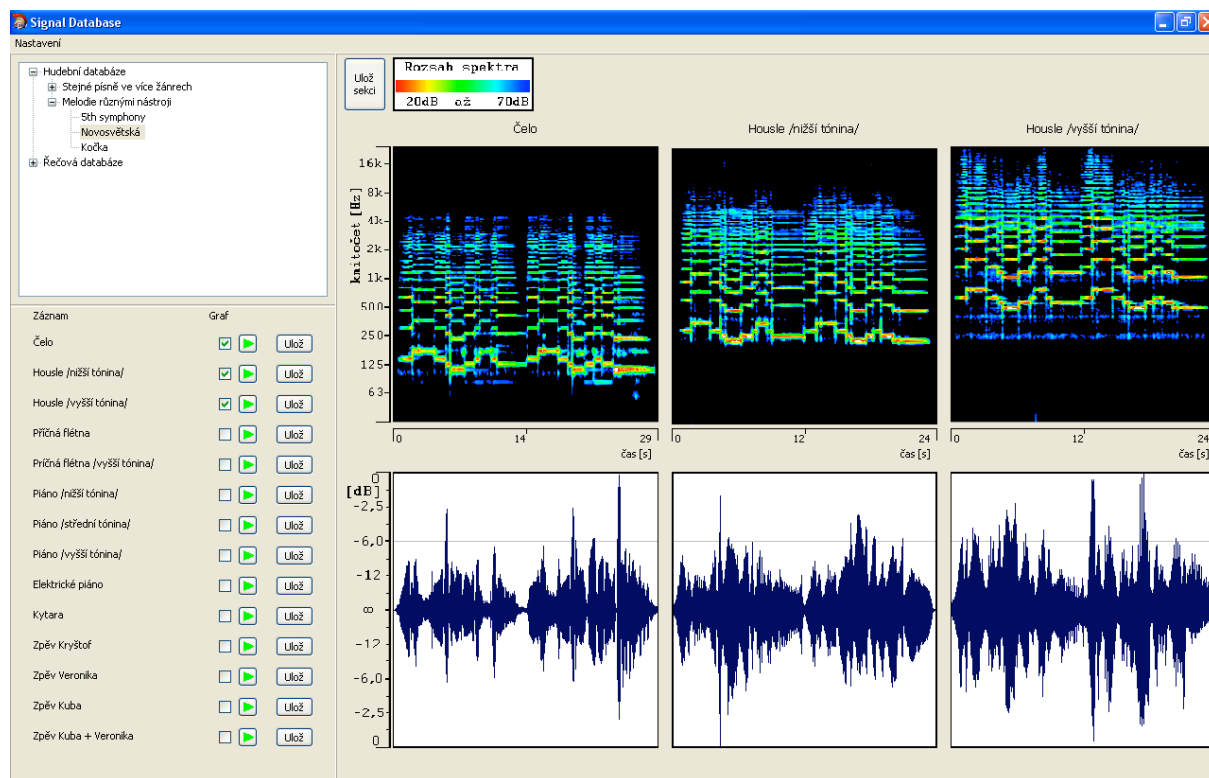
Pro další ukázkou jsem zvolil melodii ze známé básničky „kočka leze dírou, pes oknem“. Všech sedm hudebních ukázek je zahráno na housle avšak v odlišných tóninách a hlavně různými technikami.

Housle Legato -  
Housle Legato +  
Housle Legato ++  
Housle Pizzicato  
Housle Sautille

Housle Spiccato  
Housle Tremolo

vázaná hra v nízké tónině  
vázaná hra ve střední tónině  
vázaná hra ve vysoké tónině  
hra bez smyčce, drknání prsty  
zvukově působí jako velmi rychlé staccato, ale je založeno na stejném principu jako tremolo, přičemž je zde využíváno pružnosti smyčce  
přední strana smyčce jemně „poskakuje“ nad strunou  
rychlé opakování jednoho tónu nebo rychlé střídání dvou tónů vzdálených o více než sekundu položeným smyčcem

Další hudební ukázky jsem tvořil již pod hlavičkou bakalářské práce, беру je nyní jako zvuky dodané. Jedná se o melodii hlavní téma z druhé věty Dvořákovy Novosvětské hrané na housle, violoncello, příčnou flétnu, piano, elektrické klávesy, kytaru a zpěv v různých tóninách a variacích.



**Obr. 5:** Hlavní okno programu, zobrazení tří nástrojů při hře stejné melodie.

## 10.4 Stejná píseň ve více žánrech

V této části databáze se nachází hudební soubory, které reprezentuje pouze jedna píseň, ale vždy ve více hudebních verzích, žánrech, či provedeních. Máme zde několik ukázek:

### The Phantom Of The Opera

Original Soundtrack	originální nahrávka ze stejnojmenného muzikálu
Nightwish	studiová verze kapely Nightwish
Sólisté m. divadla	živé vystoupení sólistů z brněnského městského divadla

### Enter Sandman

Metallica	studiová verze kapely Metallica
Apocalyptica	studiová verze smyčcového kvarteta Apocalyptica
Apocalyptica	živě z Rockfestu



## **Nothing Else Matters**

Apocalyptica	studiová verze smyčcového kvarteta Apocalyptica
Apocalyptica	živě z Rockfestu
Lucie Silvas	vokální verze
Metallica	studiová verze kapely Metallica

## **Owner of A Lonely Heart**

Yes	originální metalová nahrávka kapely Yes
Jon Anderson	člen kapely Yes hraje akusticky /youtube.com/ <a href="http://www.youtube.com/watch?v=JdQt43KORr0">http://www.youtube.com/watch?v=JdQt43KORr0</a>
Max Graham	diskotéková „předělávka“
Alex DePue	hrané na violu /youtube.com/ <a href="http://www.youtube.com/watch?v=kIMnXyr1kf8">http://www.youtube.com/watch?v=kIMnXyr1kf8</a>

## **10.5 Řečová databáze**

V této kategorii jsou shromážděny řečové záznamy zaznamenané při určitém emočním vypětí mluvčích. Víme, že skutečný emoční stav a simulovaný emoční stav hraný hercem se v určitých detailech liší.

## **10.6 Databáze řeči pod stresem**

### **Kateřina Neumannová**

V této kategorii jsou shromážděny záznamy řeči Kateřiny Neumannové v různých situacích a tím pádem i pod různým stresem.

Tisk. konf. před závody v Liberci

Kateřina Neumannová na tiskové konferenci k závodům v Liberci:

Soubor: Katerina\_Neumannova\_tisk\_konference\_pred\_zavody\_v\_Liberci.wav

Text conference: <http://www.vlada.cz/scripts/detail.php?id=37707>

Zdroj: <http://www.vlada.cz/assets/media-centrum/tiskove-konference/tk-liberec-2008--2-7-08.mp3>

Před Libercem

Kateřina Neumannová ve studiu k závodům v Liberci (před zahájením závodů):

Soubor: Katerina\_Neumannova\_pred\_Libercem.wav

Zdroj: <http://www.ceskatelevize.cz/ivysilani/209471290170001-ms-v-klasickem-lyzovani-2009-liberec/>

Po Liberci

Kateřina Neumannová po ukončení závodů v Liberci

Soubor: Katerina\_Neumannova\_po\_Liberci.wav

Zdroj: <http://www.ceskatelevize.cz/ivysilani/209471290170041-ms-v-klasickem-lyzovani-2009-liberec/>

Ukončení Liberce

Kateřina Neumannová ukončuje závody v Liberci

Soubor: Katerina\_Neumannova\_ukonceni\_Liberce.wav

Zdroj: <http://www.ceskatelevize.cz/ivysilani/209471290170043-ms-v-klasickem-lyzovani-2009-liberec/>

Po vítězství na OH

Soubor: Katerina\_Neumannova\_po\_vitezstvi\_na\_OH\_2006.wav

Zdroj: <http://www.youtube.com/watch?v=m5HcKhoRrg8>

### **Pošta pro tebe**

V této kategorii jsou shromážděny záznamy řečových signálů prostých lidí při vysokých emočních situacích. Jedná se o zábavní televizní pořad Pošta pro tebe.

### **Dokumenty**

V této kategorii jsou shromážděny záznamy řeči autentických výpovědí lidí, kteří vzpomínají na tragické události.

Kornel Kalina

Stanislawa Rylko

Zdroj: Z pořadu Zločin jménem Katyn, vysíláno Českou televizí 11.1.2007 - 20:00

<http://www.ceskatelevize.cz/vysilani/10096326715-zlocin-jmenem-katyn/406235100021002.html>

Paní Sládková

Pan Sládek

80ti letí manželé Sládkovi přišli o střechu nad hlavou díky podvodníkovi s nemovitostmi.

Zdroj: Události ČT

### **MiniDisc**

V této kategorii jsou shromážděny záznamy řečových projevů lidí, kteří spontánně reagují na nečekané události.

## 10.7 Databáze českých mluvčích

Emoční databáze českých mluvčích je tvořena záznamy rozhlasového a televizního vysílání. Záznamy byly pořizovány z diskuzních pořadů, soutěží, reality-show, talk-show a z reportáží.

Záznamy jsou ve formátu WAV, vzorkovací kmitočet 16kHz, 16bitů, mono.

Jazyk: čeština

Druh záznamu: přirozený

Orientace v databázi:

STAV - databáze seříděná podle stavu mluvčích

Soubor: XX-STAV-Y(pozn.)

XX - označení mluvčího

STAV - stav mluvčího

Y - číslo záznamu

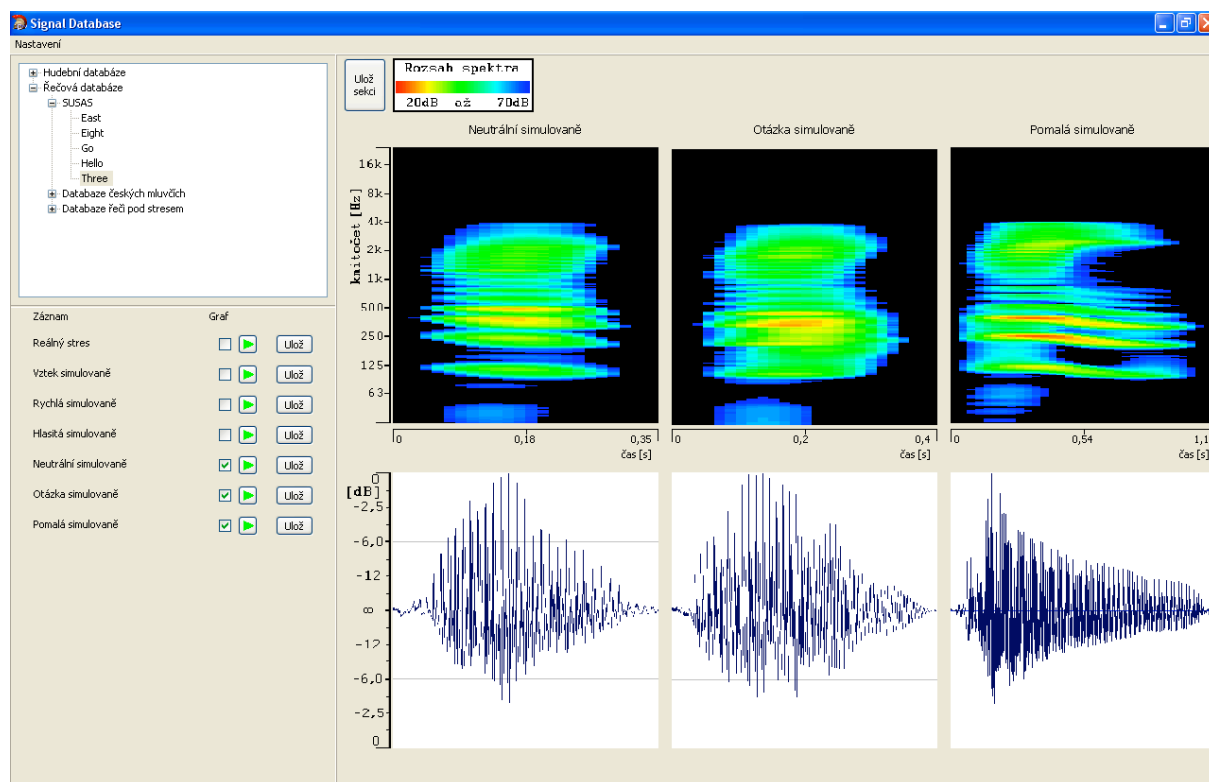
pozn. - poznámka k záznamu

Stav	Počet mluvčích	Počet záznamů	Celková délka záznamů [min:s]	Poznámka
Hněv	19	61	21:36	
Neutrální	9	11	10:53	
Překvapení	1	2	0:26	
Radost	6	7	2:04	
Smutek	9	12	8:36	
Strach	3	4	0:34	
Stres	6	51	20:40	
Únava	1	2	3:04	fyzická/psychická
Zklamání	3	13	1:11	
Žal	2	2	1:08	

**Tab. 1: Popis vlastní databáze českých mluvčích**

## 10.8 SUSAS

Zde je pro ukázkou pouze nepatrná část databáze SUSAS. Bylo vybráno pět anglických slov. Jsou zde porovnávány jak nahrávky pod reálným stresem, tak pod simulovaným. Reálný stres byl nahráván z komunikace letadel v letových potížích. Pro simulovaný stres byl vybrán mluvčí „GENERAL3“ s několika obměnami simulovaného stylu řeči: pomalá, rychlá, hlasitá, s obsahem negativních emocí – vzteku, tázací a neutrální.



**Obr. 6:** Hlavní okno programu, zobrazení jednoho slova při třech emočních vypětích.

## 10.9 Spektrogramy a časové průběhy

Spektrogramy a časové průběhy je možno zobrazit u většiny hudebních ukázek, především u těch u kterých to má smysl. Ve skupině Stejná píseň ve více žánrech je z každé písně zkopírována hlavní melodie a ta je uložena zvlášť s poznámkou /hl. melodie/. U těchto ukázek hlavní melodie lze zobrazit spektrogram a časový průběh, u celé skladby to nemá smysl. U řečových ukázek lze zobrazit časové spektrogramy a časové průběhy pouze u databáze SUSAS. U delších řečových záznamů je to nesmyslné.

## ***10.10 Jazykové rozhraní***

Jazykové rozhraní je v programu ve dvojím jazyce, a to v češtině a v angličtině. Přeloženy jsou menu, stromová struktura, názvy nástrojů, popisné osy spektrogramů, časových průběhů i legenda.

Změna jazyka na češtinu

Změna jazyka na češtinu je možná v hlavním menu, záložka Settings, možnost Change language setting.

Change Language setting to English

Change language setting to English is possible from main menu, bookmark Nastavení, option Změna jazykového nastavení.

## ***10.11 Instalace databázového serveru***

- V adresáři Signal\_Database\firebird\bin se nachází databázový server, který je třeba jednorázově spustit pomocí dávkového souboru instal\_super.bat. Databázový server pak běží jako služba operačního systému (fbserver) a spouští se automaticky po spuštění počítače. Současně se jako služba operačního systému spouští proces fbguard, který se stará o to, aby proces fbserver byl nastartován, pokud z jakéhokoli důvodu neběží.

## ***10.12 Vytvoření databáze***

- V adresáři firebird v souboru aliases.conf upravíme cestu k naší konkrétní databázi sound.fdb. Např.: sounddb = d:\Signal\_Database\firebird\sound.fdb
- Poté pomocí skriptu soundDb\_create.sql vytvoříme databázi příkazem spuštěným z příkazového řádku, např.:  
`d:\Signal_Database\firebird\bin\isql.exe -i d:\Signal_Database\soundDb_create.sql`
- Při úspěšném provedení skriptu se na obrazovce zobrazí dva řádky s textem:  
Use CONNECT or CREATE DATABASE to specify a database.

Poznámka:

Tento příkaz je vhodné použít pouze pro vytvoření databáze a tehdy, je-li nutno databázi znovu inicializovat, protože jinak hrozí ztráta obsahu celé databáze.

### 10.13 Přidávání zvukových záznamů do databáze

- V dalším kroku je třeba databázi naplnit daty, která popisují zařazení jednotlivých zvukových vzorků do kategorií, skupin a sekcí. K tomuto účelu použijeme opět příkaz isql.exe a skript nazvaný soundDb\_data.sql. Při úspěšném provedení skriptu se na obrazovce rovněž zobrazí řádek s textem:

Use CONNECT or CREATE DATABASE to specify a database.

- Uvedený skript může sloužit rovněž jako šablona pro obdobné skripty potřebné pro doplnění nových dat do databáze. Přitom je vždy nutné mít na paměti, že není možné zadat dva či více záznamů obsahujících stejný jedinečný klíč (viz řádek PRIMARY KEY v definici tabulek ve skriptu soundDb\_create.sql).
- Soubor soundDb\_data.sql můžeme upravovat v libovolném textovém editoru. Efektivně se upravuje v programu PSPad, který je možno jako freeware stáhnout ze stránky <http://www.pspad.com/cz/> a který umožňuje zvýraznění syntaxe příkazů jazyka SQL, použitého ve skriptech
- Databáze umožňuje trojí adresářové vnoření, označených jako kategorie, skupina a sekce, počet položek v jednotlivých úrovních není limitovaný.
- Každou vrstvu popisují dvě pole, první určuje český a anglický název, může vypadat například takto:

VALUES (4, 2, 'Databáze českých mluvčích', 'Database of czech speakers', 3);

- o První číslo v závorce udává jedinečný identifikátor položky v dané úrovni. Není možné zadat více záznamů se stejnou hodnotou tohoto identifikátoru.
- o Druhé číslo udává odkaz na zařazení do vyššího adresáře (v nejvyšší vrstvě pochopitelně není).
- o Třetí číslo udává pořadové číslo adresáře v daném adresáři, čili pořadí, v jakém jsou jednotlivé záznamy zobrazeny. I tato hodnota musí být jedinečná, aby bylo možné zobrazit položky ve správném pořadí.
- o V prvních apostrofech je český a v druhých anglický text, který je pak použit pro zobrazení v okně se stromovou strukturou a v seznamu skladeb.

- Druhé pole odkazuje na konkrétní adresář, může vypadat například takto:

VALUES (4, 'Databaze ceskych mluvcich');

- o Číslo udává odkaz na odpovídající položku tabulky s popisem (viz výše).
- o Název v apostrofech udává skutečný název existujícího adresáře.

- Konkrétní audio soubor popisujeme následujícím polem:

VALUES (65, 8, 'Vztek', 'Angry', 2, 'EAST1\_angry.wav');

- o První číslo v závorce udává jedinečný identifikátor záznamu.
- o Druhé číslo udává odkaz na zařazení do sekce.
- o Třetí číslo udává pořadí v adresáři pro zobrazení v programu.
- o V apostrofech jsou postupně český název, anglický název a originální název souboru včetně koncovky.

- Program se spouští souborem signalDb.exe

## **10.14 Informace**

Celý program je psán v jazyce C++ v programu CodeGear™ C++ Builder® 2007 a pracuje s databází Firebird 2.2, která je freeware. Informace o této databázi je možné získat ze stránek výrobce: <http://www.firebirdsql.org/>

## **11. Nahrávání a generování zvuků**

Živé nástroje byly nahrány do počítače programem Cubase System 4. Použil jsem zvukovou kartu M-Audio Fast Track Pro a dvojici studiových kondenzátorových mikrofónů Rode NT1-A. Zvolil jsem standardní nahrávací metodu jednoho nástroje dvěma mikrofony. Umístění mikrofónů membránami vertikálně proti sobě, osově jemně vychylené. Nebyla použita žádná ekvalizace, změna dynamiky, či jiný proces. Pouze normalizace na hodnotu RMS = -15,55dB a Peak = -0,22dB. Všechny živé zvuky jsou nahrány v tempu 133 BPM.

Virtuální nástroje byly generovány v programu Cubase System 4 virtuálními midi nástroji firmy Steinberg. Normalizace byla provedena v programu Steinberg Wavelab 6 na hodnotu RMS = -15,55dB a Peak = -0,22dB. Všechny virtuální zvuky jsou nahrány opět v tempu 133 BPM. Obrázky zjednodušených notových partitur byly generovány programem Cubase System 4.

## **12. Formát souborů v databázi**

Všechny přiložené zvukové soubory jsou ve standardním zvukovém formátu \*.wav při vzorkovací frekvenci 44,1kHz, bitové hloubce 16bitů, PCM, hudební soubory stereo, řečové mono.

Při převodu souborů databáze SUSAS bylo zachováno originální rozlišení. Vzorkovací frekvence 8kHz, bitová hloubka 16bitů, mono, PCM.

## 13. Všeobecné audio formáty

### 13.1 *Lidské vnímání zvuku*

Lidské vnímání zvuku je velice složitý proces, závislý na mnoha faktorech, pro který zatím nebyla vytvořena uspokojivá teorie.

#### Frekvenční rozsah

Frekvenční rozsah zvuku, který většina lidí vnímá, začíná kolem 20 Hz a dosahuje ke 20 kHz. S rostoucím věkem horní hranice výrazně klesá. Nejvýznamnější rozsah je 2–4 kHz, který je nejdůležitější pro srozumitelnost řeči a na něj je lidské ucho nejcitlivější. Nejvyšší informační hodnota řeči je přenášena v pásmu 0,5–2 kHz.

#### Dynamický rozsah

Dynamický rozsah lidského ucha (rozdíl mezi nejhlasitějším a nejtisším vnímatelným zvukem) je uprostřed slyšitelného frekvenčního pásma asi 120 dB. Na okrajích pásma je mnohem menší.

#### Rozlišování frekvence

Schopnost rozlišit frekvence tónů se u každého člověka liší a je frekvenčně závislá. Uprostřed slyšitelného frekvenčního pásma za ideálních podmínek lze rozlišit změnu frekvence o několik centů. Na okrajích pásma je rozlišovací schopnost výrazně nižší.

#### Frekvenční maskování

Schopnost odlišit dva frekvenčně blízké tóny je ovlivněna frekvenčním maskováním. Pokud znějí dva tóny současně, může jeden z nich potlačit slyšitelnost toho druhého. Tato neschopnost slyšet oba současné tóny se nazývá frekvenční maskování. Maximální úroveň maskovaného signálu je závislá na frekvenční vzdálenosti a úrovni maskujícího signálu. Maskovací schopnost je též závislá na frekvenci maskujícího tónu. Vnímání tónů s blízkými frekvencemi je ovlivněno šířkou kritického pásma. To má na nejnižších kmitočtech velikost kolem 100 Hz, zatímco na nejvyšších kmitočtech dosahuje až 4 kHz.

#### Časové maskování

Pokud po hlasitém tónu následuje stejný tón s menší hlasitostí, je jeho vnímání potlačeno. Potlačen může být i tichý tón předcházející maskovacímu tónu.

Ztrátové formáty fungují tak, že je část informace, která není pro lidské ucho podstatná úplně vypuštěna. Jedná se o tu část dat, která je nejméně důležitá pro poslech a běžný posluchač ji v nahrávce ani neslyší. Kdežto u bezztrátových formátů se žádná informace nevypouští. [2]



## **13.2 Vlastní audio formáty**

### **13.2.1 MP3**

MP3 (MPEG-1 Layer III) je formát ztrátové komprese zvukových souborů, založený na kompresním algoritmu MPEG (Motion Picture Experts Group). Při zachování poměrně vysoké kvality umožňuje zmenšit velikost hudebních souborů v CD kvalitě přibližně na desetinu, u mluveného slova však dává výrazně horší výsledky.

Formát MP3 se stal oblíbeným při uchovávání a přehrávání hudby na počítačích, vyrábí se stolní a přenosné přehrávače tohoto formátu.

O vývoj formátu MP3 se zasloužil německý vědec Karlheinz Brandenburg, ředitel pobočky Fraunhoferova ústavu pro mediální komunikaci v Ilmenau, a jeho vědecký tým.

MP3 se snaží odstranit redundanci zvukového signálu na základě psychoakustického modelu. Tedy ze vstupního signálu se odeberou informace, jež člověk neslyší, nebo si je neuvědomuje. Využívá se principů časového a frekvenčního maskování. Komprese zvuku podle standardu MPEG-1 obsahuje 3 vrstvy, jež se liší kvalitou a obtížností implementace.

Slabiny:

Při kompresi mluveného slova jsou výsledky výrazně horší. Popsané maskování a potlačování tónů způsobuje, že u mluveného slova může být ve slově potlačena počáteční nebo koncová slabika. Mohou být také zkracovány pauzy mezi jednotlivými slovy. To působí u mluveného slova značně rušivě. Pro kompresi hlasu jsou vhodné jiné metody např. AMR, G.729, OGG Vorbis, nebo Speex.

Výsledná kvalita ovšem závisí na zvoleném datovém toku. [2]

### **13.2.2 AAC**

Stejně jako MP3 i formát AAC - Advanced Audio Coding je součástí MPEG-2 audio standardu. Autorem AAC je několik výzkumných firem a prodej licence pro jeho další používání zajišťuje firma Dolby Laboratories. Po technické stránce nabízí AAC lepší kompresi při zachování srovnatelné kvality s MP3 (výsledkem jsou menší komprimované audio soubory), větší počet audio kanálů a širší nabídku dostupných vzorkovacích frekvencí pro zajištění lepšího výsledného zvuku, než nabízí formát MP3. Přestože formát AAC zatím nebyl masově použit ani rozšířen, posloužil jako základ několika patentovaných zabezpečených audio formátů a bude rovněž součástí standardu MPEG-4. [10]

### **13.2.3 ATRAC3**

Formát Adaptive Transform Acoustic Coding 3 (ATRAC3) je majetkem firmy Sony a představuje vylepšenou verzi formátu ATRAC, který je používán v přehrávačích MiniDisc. ATRAC3 nabízí stejnou kompresi i zvukovou kvalitu jako MP3. V současné době používá ATRAC3 pouze firma Sony ve svých přenosných přehrávačích (například VAIO Music Clip), ale o jeho licencování projeví zájem několik dalších firem, které se zabývají vývojem a výrobou přenosných audio přehrávačů. Za všechny jmenujme alespoň firmu e.Digital. Podle vyjádření firmy Sony by měl být formát ATRAC3 SDMI kompatibilní. [10]

### **13.2.4 ePAC**

Enhanced Perceptual Audio Coder (ePAC) firmy Lucent Technologies je podle tiskové zprávy při kompresním poměru 11:1 k nerozeznání od originálního CD. Tento kompresní poměr jej staví téměř na stejnou úroveň jako formát MP3. Přestože ePAC nepatří mezi příliš populární kompresní audio formáty, o jeho licencování projevil zájem několik předních firem vyrábějících přenosné MP3 přehrávače. Za zmínku rovněž stojí skutečnost, že firma Maycom (výrobce oblíbeného přehrávače I-Jam) licencovala od partnera Lucent Technologies firmy eDigital kompletní design ePAC přehrávače. Formát ePAC je SDMI kompatibilní a vše nasvědčuje tomu, že o něm ještě uslyšíme. [10]

### **13.2.5 Liquid Audio**

Liquid Audio patří mezi formáty postavené na technologii AAC a využívá bezpečnostního kódování pro ochranu zabezpečených audio souborů proti ilegálnímu kopírování. Liquid Audio patří mezi relativně hodně rozšířené kompresní formáty. V současné době existuje několik přenosných přehrávačů od předních světových výrobců, které formát Liquid Audio podporují. Rovněž po dohodě s AOL se podpora tohoto formátu časem objeví ve Winampu. Celá řada předních populárních i nezávislých hudebníků a kapel nabízí svoje nahrávky právě ve formátu Liquid Audio. Liquid Audio je SDMI kompatibilní. [10]

### **13.2.6 Real Audio**

RealNetworks byli první firmou, která přišla s kompresním formátem pro streaming audio a video dat. Ještě nedávno jejich formát Real Audio zaujímal téměř monopolní postavení na trhu streamovaných multimedií. Nyní soupeří o dominantní postavení s formátem Windows Media firmy Microsoft. Formát Real Audio je po zvukové stránce při vyšším datovém toku (bitrate) srovnatelný s MP3, ale jeho hlavní využití stále spočívá ve streamování audio dat nižší kvality. [10]

### **13.2.7 VQF (Twin VQ)**

VQF je zkratkou pro Transform-domain Weighted Interleave Vector Quantization formát, který se také občas nazývá TwinVQ. Tento formát vyvinula firma Nippon Telegraph and Telephone (NTT) a nabízí lepší kompresi i lepší zvukovou kvalitu než formát MP3. Přestože firma Yamaha uvedla na trh freeware VQF enkodér i přehrávač, formát si příliš příznivců nezískal. Hlavním důvodem je pravděpodobně podstatně delší doba potřebná pro komprimaci audio souboru a větší hardwarové nároky při jeho přehrávání v porovnání s formátem MP3. Formát VQF bude součástí standardu MPEG-4. [10]

### 13.2.8 Windows Media Audio (WMA)

Formát Windows Media Audio (WMA) je součástí programového vybavení Windows Media firmy Microsoft. Podle prohlášení firmy Microsoft nabízí zvukovou kvalitu srovnatelnou s audio CD při datovém toku 64kbps, což představuje přibližně dvakrát účinnější kompresi oproti MP3. Tuto skutečnost potvrdil nezávislý výzkum ZDLabs (jako MP3 kodek byl použitý software firmy Xing, který nenabízí nejlepší kvalitu výsledných MP3 souborů). WMA je zabezpečený SDMI kompatibilní formát. Krátce po svém uvedení na trh v loňském roce byl však prolomen. Ve spojení s ASF - Advanced Streaming Format může být WMA streamován. Díky tomu se formát Windows Media stal přímým konkurentem zavedeného formátu RealAudio a velmi rychle získává na oblibě. V současné době WMA formát podporuje celá řada přenosných hardwarových i softwarových přehrávačů. Formát WMA je jednoznačně hlavním kandidátem na přední místo mezi kompresními audio formáty blízké budoucnosti. [10]

### 13.2.9 Vorbis (kodek)

Vorbis je svobodný a open source ztrátový audio kodek, který se měl stát náhradou formátu MP3. Jeho vývoj je zastřešen organizací Xiph.Org Foundation. Nejčastěji bývá uložen v kontejneru Ogg a v tomto spojení je nazýván Ogg Vorbis. Kodek Vorbis se stal prvním Xiph kodekem zralým pro všeobecné použití, a po jistou dobu v podstatě jediným. Tento kodek byl v minulosti bohužel často zaměňován s kontejnerovým formátem OGG.

#### Kvalita

Po stránce kvality je zvukový kodek přibližně na stejné úrovni jako konkurenční formát Microsoftu Windows Media Audio. Oba jsou kvalitnější než dnes již zastaralý, ale stále nejpoužívanější formát MP3.

Přesné srovnání je bohužel téměř nemožné, už kvůli principu ztrátové komprese. Protože každé lidské ucho má jiné akustické vnímání, není možné jednoznačně určit, který ze dvou kvalitativně podobných akustických modelů je lepší nebo kvalitnější.

#### Použití

Kromě použití jako kompresoru hudebních souborů u domácích uživatelů se začal poměrně často používat výrobci her pro kompresi zvukových (někdy i obrazových) dat ve hrách obsažených. [2]

### 13.2.10 AMR (kodek)

AMR (anglicky Adaptive Multi-Rate compression) je metoda komprese zvuku používaná a určená především pro řeč. AMR je kodek používaný v UMTS sítích ale je ho možno použít i GSM sítích dále se používá zejména v elektronických diktafonech a mobilních telefonech, kde při kompresi mluveného slova dosahuje výborných výsledků. AMR má celou řadu režimů a je kompatibilní s GSM-EFR, PDC a TMDA-EFR.

Horší kvalitu má záznam hudby, pro který je vhodnější např. formát MP3 či WMA. AMR se také používá pro kompresi zvuku v MMS zprávách a také v systému IMS je základním kodekem. AMR je součástí standardu 3GPP. Označení AMR se užívá rovněž jako přípona souborů touto metodou komprimovaných (\*.amr). Vzorkovací frekvence je 8 kHz při 13 bitech, každých 20 ms je možná změna bitového toku, proto název Adaptive Multi-Rate. Přes svoje značné rozšíření není samotná existence formátu mezi jeho uživateli příliš známa.

AMR kodek umožňuje 14 módů kódování, 8 je pro plnou rychlost (full rate) a 6 pro poloviční rychlost (half rate).

Mód	Rychlost	Kanál	Kompatibilní s
AMR_12.20	12,20 kbit/s	FR	ETSI GSM Enhanced Full Rate
AMR_10.20	10,20 kbit/s	FR	
AMR_7.95	7,95 kbit/s	FR/HR	
AMR_7.40	7,40 kbit/s	FR/HR	TIA/EIA IS-641 TDMA Enhanced Full Rate
AMR_6.70	6,70 kbit/s	FR/HR	ARIB 6,7 kbit/s Enhanced Full Rate
AMR_5.90	5,90 kbit/s	FR/HR	
AMR_5.15	5,15 kbit/s	FR/HR	
AMR_4.75	4,75 kbit/s	FR/HR	
AMR_SID	1,80 kbit/s	FR/HR	

#### AMR-WB

AMR-WB neboli AMR WideBand používá dvojnásobnou vzorkovací frekvenci tedy 16 kHz a umožňuje vyšší datové toky a tím vyšší kvalitu. AMR-WB poskytuje tedy věrnější přenos a je s ním počítáno třeba pro mobilní televizi nebo rádio. Také je možné použití pro kvalitnější telefonní hovor nebo video či telekonference.

#### AMR WB+

AMR WB+ tedy AMR WideBand plus, někdy také Stereo over AMR, umožňuje stejně jako AMR-WB použití vzorkovací frekvence 16 kHz nebo 8 kHz a přidává druhý kanál pro přenos stereofonního signálu. [2]

### 13.2.11 Apple Lossless

Apple Lossless (také známý jako Apple Lossless Encoder, ALE případně Apple Lossless Audio Codec, ALAC) je audio kodek, vyvinutý společností Apple Computer pro bezztrátovou kompresi hudby.

Pro ukládání hudby kódované pomocí Apple Lossless je použit kontejner MP4; soubory mají příponu .m4a. Apple Lossless není založen na AAC, ale používá lineární predikci podobně jako ostatní bezztrátové audio kodeky, například FLAC a Shorten. Není použita žádná metoda DRM, ale díky povaze kontejneru je možné DRM aplikovat na ALAC stejně jako u ostatních souborů využívajících QuickTime kontejnery.

Apple tvrdí, že audio soubory komprimované pomocí ALAC budou potřebovat přibližně polovinu velikosti nekomprimovaných dat. Testy však bylo zjištěno, že soubory průměrně dosahují 60 % původní velikosti, podobně jako ostatní bezztrátové formáty. Avšak na rozdíl

od ostatních formátů, Apple Lossless není náročný na dekompresi a je tak použitelný i v přístrojích s omezeným zdrojem energie jako například iPod.

Apple Lossless byl představen 28. dubna 2004 jako součást QuickTime verze 6.5.1, je tudíž použitelný v iTunes verze 4.5. Kodek je také používán v AirPort Extreme.

David Hammerton a Cody Brocious analyzovali a dekodovali tento formát bez jakékoli dokumentace. 5. března 2005 Hammerton publikoval jednoduchý open source dekodér, napsaný v programovacím jazyku C.

Ve starších přehrávačích společnosti Apple lze tento formát přehrát s firmwarem Rockbox, který je poskytován zdarma. [2]

### **13.2.12 G.711**

G.711 je nejběžnější, nejrozšířenější a nejjednodušší kodek používaný v telekomunikacích. Bohužel pro svou jednoduchost je také nejméně úsporný na přenosovou rychlost. Vzorkovací frekvence u tohoto kodeku je 8 kHz a rozlišení 8 bitů, z toho vyplývá přenosová rychlost 64 kbit/s. Kodek G.711 je použit u v ISDN a PCM.

Pro kódování signálu se používá logaritmická komprese, kdy je dvanácti- či třinácti bitový signál převáděn na osmibitový signál. Logaritmická komprese hovorového signálu je v Evropě a Austrálii známá jako A-law, v severní Americe a Japonsku jako  $\mu$ -law. V severní Americe a Japonsku se používá vyšší komprese, protože v tamních telefonních sítích je používáno pouze 7 bitů pro přenos hlasu a osmý bit se používá na signalizaci. V Evropě se používá pro přenos hlasu všech 8 bitů a signalizace je přenášena samostatným kanálem. (Například pomocí signalizace č.7, CAS či jinou signalizací.) V Evropě a Austrálii má tedy přenos hlasu v digitální telefonní síti vyšší kvalitu než stejný přenos pomocí telefonní sítě v Severní Americe či v Japonsku. Na rozhraní těchto sítí musí proto dojít k překódování. Jakým způsobem je hlas kódován, tedy informace o použitém kodeku, jsou přenášeny pomocí signalizace. [2]

### **13.2.13 G.729**

G.729 je úsporný hlasový kodek, který umožňuje efektivněji využít limitované pásmo pro přenos hlasu po počítačové síti (IP telefonie). Běžně je pro hovor vyčleněno pásmo o šířce 64 kbit/s např. kodek G.711 (8 kHz, 8 bitů, ISDN a PCM tedy standardní digitální telefonní síť). Kodek G.729 umožňuje toto pásmo snížit na šířku 8 kbit/s. [2]

### **13.2.14 MOD**

MOD je oblíbený zvukový formát pro uskladnění hudby na počítačích a je prvním originálním typem modulového souboru. Tento formát byl původně vytvořen tak, aby byl

snadno hratelný na hardwaru domácího počítače Amiga. Díky tomu měl jeho pomalý 16/32-bitový procesor Motorola 68000 jen velmi málo práce.

Byl generován prvním hudebním sekvencerem SoundTracker (známým také jako ProTracker nebo FastTracker). V pozdějších variantách formátu používal až 32 hlasů. Nástroje mají jednoduché úrovně hlasitosti. Vzorky a nástroje korespondují v poměru jedna ku jedné. Originální SoundTracker měl jen 15 nástrojů, ale další sekvencery až 31 nástrojů.

Minimální časový rámec MOD souboru je 0,02 sekundy, protože originální software využíval VSync monitoru, který běžel v 50 Hz časovacím cyklu. Každá stopa měla pro sebe až 64 řádek, každá řádka trvala 0,02 sekundy času při běžném nastavení rychlosti. [2]

### 13.2.15 Speex

Speex je svobodný zvukový kodek šířený pod licenci BSD, který slouží ke kompresi lidské řeči. Je součástí projektu ogg. Jeho specializace mu umožňuje dosahovat velmi dobrých kompresních poměrů při zachování srozumitelnosti. Není primárně určen (jako ostatní podobné kodeky) na kompresi telefonních hovorů, ale na VoIP (Voice over IP).

Speex byl navržen pro použití se třemi vzorkovacími frekvencemi 8 kHz, 16 kHz, a 32 kHz (narrowband, wideband a ultra-wideband). Kvalita se nastavuje jako parametr od 0 do 10. CBR se nastavuje celočíselným parametrem, VBR reálným parametrem.

Speex má několik zajímavých postupů, které jsou speciálně určeny pro kompresi lidské řeči:

- \* Voice Activity Detection (VAD) - česky detekce aktivity hlasu. Tato technologie má za úkol rozpoznat, zda řečník právě mluví, nebo dělá přestávku (mezi slovy, větami), nebo prostě jen čeká. V tom případě je produkováno jen velmi malé množství bitů, které stačí na generování šumu.

- \* Discontinuous Transmission (DTX) - rozšíření předchozí technologie, která detekuje delší odmlčení. V tom případě úplně zastavuje přenos dat.

- \* Perceptual enhancement - Snaží se zeslabit šum vytvořený kódováním/dekódováním. Objektivně (tedy měřením) tato technologie vzdaluje výstupní zvuk od originálu, ale subjektivně je zvuk lepší. [2]

### 13.2.16 Flac

Flac (Free Lossless audio codec) je kodek pro formát, který využívá bezztrátovou kompresi, takže naše hudba je uchována v naprosto dokonalé kvalitě. Nabízí 8 stupňů komprese, při největším stupni je soubor zmenšen na 60-70% své původní velikosti, jedna písnička má tedy velikost mezi 20 a 30 megabajty. Pokud chcete tento formát přehrát na počítači, je nutné si nejprve doinstalovat kodek. Co se týká podpory v přenosných přehrávačích, ta je zatím slabá, flac podporuje pouze pár zařízení, včetně populárního iPodu. Stejná situace je i u hifi věží a autorádií. Vhodný pro: vysoce kvalitní přehrávání hudby bez jakýchkoli ztrát, pro kvalitní audio sestavu a sluchátka k PC.[2]

### 13.2.17 WAV

WAV (nebo také WAVE) je zkratka a běžně používaná přípona pro Waveform audio format. Tento zvukový formát vytvořily firmy IBM a Microsoft pro ukládání zvuku na PC. Je to speciální varianta obecnějšího formátu RIFF.

Přestože je možné ukládat do WAV souboru zvuk komprimovaně, například pomocí GSM komprese, ADPCM,  $\mu$ -Law či v MP3, většinou se používá nekomprimovaný zvuk v pulzně kódové modulaci. Stejným způsobem je uložen zvuk na Audio CD, což umožňuje snadný převod mezi těmito formáty.

Protože PCM je bezztrátový formát, používá se nejčastěji WAV při zpracování zvuku. [11]

### 13.2.18 Aiff

Obdobně na platformě Apple je využíván formát AIFF (Audio Interchange File Format), jehož soubory mají příponu .aif. V dřívějších dobách této platformě vévodil formát Sound Designer. Byl to vlastně předchůdce AIFF souborů, který byl využíván systémem Digidesign ProTools. V dnešní době je v tomto systému možno pracovat s formáty AIFF i Wave a to na obou platformách, tedy na PC i na Macu. Také většina ostatních profesionálních DAW aplikací podporuje oba typy dat alespoň na úrovni importu a exportu. [2]

### 13.2.19 DVD-Audio

DVD-Audio (obvyklá zkratka: DVD-A) je digitální formát, dodávající vysoce kvalitní reprodukci zvuku, obsaženou na médiích typu DVD. DVD-Audio neobsahuje video, nemělo by tedy být zaměňováno s video DVD, které obsahuje hudební videa. První disky tohoto formátu vstoupily na trh v roce 2000. Data na těchto discích jsou ve formátu \*.war, stejně jako Super Audio CD (SACD), což je další formát pro reprodukci vysoce kvalitního zvuku.

#### Specifikace zvuku

DVD-Audio nabízí mnoho možných konfigurací zvukových kanálů, od jednobanálních mono po 5.1 prostorový zvuk v různých vzorkovacích kmitočtech a vzorových sazbách. Ve srovnání s CD umožňuje větší kapacita DVD média zahrnout následující:

- \* podstatně více hudby (s ohledem na celkový čas a množství písní)
- \* daleko vyšší zvuková kvalita
- \* další kanály pro reprodukci prostorového zvuku

Audio na DVD-Audio disku může být uloženo v mnoha rozmanitých přenosových rychlostech, vzorkovacích kmitočtech a kanálových kombinacích.

Různé přenosové rychlosti, vzorkovací kmitočty a kanálové kombinace mohou být použity na jednom disku. Například DVD-Audio disk může obsahovat 96 kHz / 24-bit 5.1-prostorový zvuk, stejně jako 192 kHz / 24-bit stereo audio stopu. Audio kanály mohou být také rozděleny na dvě skupiny a uloženy pod různými formáty. Například přední reproduktory mohou

přehrávat 96 kHz / 24-bit 5.1-prostorový zvuk, zatímco postranní reproduktory přehrávají zvuk s parametry 48 kHz / 20-bit.

#### Kompatibilita s přehrávači

Uvedení DVD-Audio formátu vyžadovalo zpětnou kompatibilitu s již existujícími DVD-Video přehrávači. Proto většina DVD-Audio disků obsahuje minimálně audio stopu formátu Dolby Digital 5.1. Některé disky zahrnují zvuk formátu Dolby Digital 2.0, některé dokonce DTS 96/24 5.1.

Protože je DVD-Audio formát členem DVD rodiny, jednotlivé disky mohou obsahovat více vrstev, popřípadě dvě strany, obsahující audio data. Některé DVD tituly obsahují DVD-Audio formát na jedné straně a na druhé straně obsahují klasický formát DVD-Video, příkladem jsou disky DVD-10: dvě strany, jedna vrstva na každé straně.

Existují softwarové přehrávače, které podporují přehrávání DVD-Audia. Příkladem těchto přehrávačů jsou WinDVD a PowerDVD.

#### Kvalita zvuku

Z čistě technického hlediska je kvalita DVD-Audio disků podstatně vyšší, než je standard, definovaný v Red book (Červená kniha) pro CD audio. DVD-Audio podporuje až 24 bitovou hloubku a vzorkovací frekvenci až 192 kHz (CD audio 16-bit, 44.1kHz).

Tři z hlavních distributorů hudby - Universal Music, EMI a Warner Bros Records a několik menších organizací hudebního průmyslu dále vydává hudbu ve formátu DVD-Audio, ale poměr oproti prodaným CD audio nosičům je minimální.

Nové tituly jsou většinou uvolňovány v následujících formátech

- \* klasická audio CD
- \* ve standardním DVD-Video formátu
- \* HDAD - formát DVD-Video na jedné straně a na druhé formát DVD-Audio
- \* svazek CD/DVD - obsahuje album ve formátu CD a DVD-Audio [2]

### 13.2.20 SACD

Super Audio Compact Disc (SACD) je zvukový nosič vizuálně stejný jako klasický kompaktní disk. Nabízí však mnohem lepší zvukový komfort. Klasické CD nabízí frekvenční rozpětí do 20 kHz, SACD dokáže věrně reprodukovat zvuk až do 100 kHz. Výhodou nosiče SACD je, že dokáže navíc reprodukovat i prostorový zvuk a díky dvěma vrstvám určeným pro uložení záznamu může nést současně i stejnou stopu jako klasické CD - takové disky pak mohou být přehrávány na jakémkoli CD přehrávači nebo CD mechanice a jsou označovány jako hybridní. Pro naprosto věrné přehrávání záznamu SACD je nutné vlastnit přehrávač podporující funkce SACD a samozřejmě i kvalitní Hi-Fi zařízení. SACD v době svého vzniku mělo nejen nabídnout vynikající zvuk hudebním posluchačům, ale částečně i vyřešit propady hudebního trhu, zejména vzhledem k nemožnosti pořizování identických pirátských kopií. Nosič však v masovém měřítku neuspěl a nyní se stal spíše záležitostí pro náročné



posluchače. Množství titulů vydávaných na SACD navíc není nikterak velké. I proto je celý náklad řady titulů dávno rozebrán.

Formát SACD byl poprvé představen v roce 1999, společně ho vyvíjely firmy Sony a Philips — tedy společnosti, které společně vyvinuly i klasický CD disk. SACD vede válku s formátem DVD-Audio, ale ani jednomu formátu se zatím nepodařilo nahradit klasické CD disky. [2]

### 13.2.21 Dolby Digital

Dolby Digital (původní označení AC-3) je označení digitální ztrátové komprese zvuku, vyvinuté společností Dolby Laboratories roku 1991. Je přímým následníkem formátů Dolby Stereo, resp. Dolby Surround a konkurentem formátů DTS a SDDS. Nejčastěji se vyskytuje v konfiguraci 5.1 kanálů (pět hlavních plnorozsahových kanálů a jeden nízkofrekvenční „basový“ LFE kanál využívaný subwooferem při explozích a podobně), ale podporováno je i stereo, mono, Dolby Surround a některé další konfigurace. Plnorozsahové kanály mají rozsah od 20 Hz do 20 kHz, LFE kanál má rozsah do 120 Hz. Původně se tento formát používal v kinech (zde vždy v konfiguraci 5.1, od roku 1998 i v konfiguraci Dolby Digital Surround EX), později se začal používat i na laserdiscích, v digitálním televizním vysílání a je to i základní formát pro kódování zvuku na DVD. Dá se nicméně použít i jako součást videosouborů AVI, OGM, nebo MKV. Zvukové soubory formátu Dolby Digital mají typicky příponu ac3.

Zleva doprava:

SDDS, Dolby Digital, Dolby Stereo, DTS

Na filmovém pásu je zaznamenán v prostoru mezi perforačními otvory v levé části pásu. Zálohován je analogovým formátem Dolby Stereo.

Datové toky

Audiostandard AC-3 může mít maximální datový tok 640 kbit/s. U DVD Video má maximální datový tok 448 kbit/s, přičemž nejčastějšími datovými toky na DVD-Video jsou 192 nebo 224 kbit/s pro konfiguraci 2.0 a 2.0 Surround, 384 nebo 448 kbit/s pro konfigurace 5.0 a 5.1, a 448 kbit/s pro konfiguraci 5.1 EX. Maximální datový tok při použití u digitálního televizního vysílání je 448 kbit/s. Maximálním datovým tokem při použití na ATSC je 384 kbit/s.

Možné konfigurace

- \* Dolby Digital 1.0 – mono
- \* Dolby Digital 2.0 – obsahuje 2 zvukové kanály a může být buď stereo nebo mono.
- \* Dolby Digital 2.0 Surround – Dolby Surround, na rozdíl od analogově zaznamenaného zvuku ve formátu Dolby Surround však obsahuje navíc informaci o tom, že se jedná o Dolby Surround.
- \* Dolby Digital 4.0 – nepříliš častá varianta, obsahující 4 samostatné kanály (levý, střední, pravý a zadní).
- \* Dolby Digital 5.0 – nepříliš častá varianta (obsahuje ji například DVD Vodní svět), obsahující 5 samostatných kanálů (přední levý, střední, přední pravý, zadní levý a zadní pravý) bez basového LFE kanálu.

\* Dolby Digital 5.1 – nejčastější varianta, obsahující 5 hlavních plnorozsahových kanálů (přední levý, střední, přední pravý, zadní levý a zadní pravý) a jeden LFE kanál. Prvním filmem v tomto formátu byl Batman se vrací, z roku 1992.

#### Dolby Digital 5.1 EX

Dolby Digital 5.1 EX, často označovaný jen jako Dolby Digital EX, v kinech jako Dolby Digital Surround EX, je nejnovější variantou vytvořenou v říjnu 1998, ve spolupráci se společností Lucasfilm THX, a poprvé uvedenou ve filmu Star Wars: Epizoda I – Skrytá hrozba. Vedle klasických 5.1 kanálů obsahuje navíc ještě zadní střední kanál, který je maticově zakódovaný v zadním levém a v zadním pravém kanálu. Tento formát lze v domácích podmínkách přehrávat buď v konfiguraci 6.1 nebo v konfiguraci 7.1 (zadní střední kanál přehrávají místo jednoho reproduktoru rovnou 2 reproduktory). První DVD filmy, které byly zaznamenány s tímto zvukem, ještě v hlavičce AC-3 streamu neobsahovaly digitální informaci o tom, že tento maticově zakódovaný zadní kanál obsahují. Novější DVD filmy už v hlavičce svého AC-3 streamu tuto informaci obsahují, tudíž je tento kanál na patřičně vybavených dekodérech dekódován automaticky.

#### Dolby Digital Plus

Dolby Digital Plus je vylepšený kódovací systém založený na formátu Dolby Digital. Nabízí vyšší datové toky (až do 3 Mbit/s), podporuje více zvukových kanálů (až 13.1), vylepšené kódovací techniky redukující kompresní artefakty, a zpětnou kompatibilitu s existujícím AC-3 hardwarem. Od května 2005 se o něm uvažuje jako o jednom z možných audioformátů pro HD-DVD. [2]

## 14. Česká dokumentace pro databázi SUSAS

Jako zvláštní část práce jsem vytvořil překlad anglické dokumentace k databázi SUSAS. Je brána jako příloha k práci. Viz příloha č.1.

## 15. Převod souborů databáze SUSAS

### 15.1 *Sph convert*

Originální soubory použité databáze SUSAS se nacházely v nestandardním formátu \*.sph. Pro zajištění větší kompatibility, přístupnosti a použitelnosti jsem všechny soubory převedl do standardního zvukového formátu \*.wav a to následujícím programem „Sph convert v2.1“ pracujícím v příkazovém řádku:

[ftp://ftp.ldc.upenn.edu/pub/ldc/misc\\_sw/sph\\_convert\\_v2\\_1.zip](ftp://ftp.ldc.upenn.edu/pub/ldc/misc_sw/sph_convert_v2_1.zip)

Přesný návod a zdrojové informace zde:

[http://library.rice.edu/services/digital\\_media\\_center/online\\_guides/working-with-linguistic-resources/converting-sph-audio-files-to-wav](http://library.rice.edu/services/digital_media_center/online_guides/working-with-linguistic-resources/converting-sph-audio-files-to-wav)

## 15.2 Sound Converter

K převodu je možné použít i program Sound Converter, který je jednodušší, kompaktnější a podporuje stromovou strukturu adresářů. Zadarmo pracuje se soubory do 500kB pod OSX nebo 14 dní pod MS Windows. Plná licence programu stojí \$14.

Program Sound Converter lze stáhnout z následující stránky:

<http://www.dekorte.com/projects/shareware/SoundConverter/>

Podporované typy souborů pro operační systém OSX :

Vstupní formáty:

4xm, 8svx, ac3, aif, aifc, aiff, al, ape, au, auto, avi, avr, caf, cdr, cvs, dat, dv, fap, ffm, flac, gsm, h263, h264, hcom, ircam, kmp, ksf, la, lu, m4v, mat, mat4, mat5, maud, mov, mp2, mp3, mpc, mpeg, mpegts, nist, nul, ogg, paf, prc, pvf, raw, redir, rtp, rtsp, sl6be, sl6le, s8, sb, scm, sd2, sdp, sf, sffd, shn, sl, smp, snd, sndt, sph, svx, sw, tta, txw, u16be, u16le, u8, ub, ul, uw, vms, voc, vorbis, vox, w64, wav, wve, xi

Výstupní formáty:

8svx, aac, ac3, aif, aifc, aiff, al, ape, au, auto, avr, caf, cdr, crc, cvs, dat, dv, fap, ffm, flac, gsm, h263, h264, hcom, ircam, la, lu, m4v, mat, mat4, mat5, maud, mp2, mp3, mp4, mpeg, mpegts, nist, nul, ogg, paf, prc, pvf, raw, rtp, sl6be, sl6le, s8, sb, sd2, sf, shn, sl, smp, snd, sndt, sph, svx, sw, swf, tta, txw, u16be, u16le, u8, ub, ul, uw, vcd, vms, vob, voc, vorbis, vox, w64, wav, wve, xi

Podporované typy souborů pro operační systém MS Windows :

Vstupní formáty:

aac, aif, aiff, amr, au, flac, gsm, m4a, mp3, ogg, raw, vox, wav, wma, 4xm, 8svx, aac, ac3, aif, aifc, aiff, al, amr, ape, asf, au, auto, avi, avr, cdr, cvs, dat, dv, fap, ffm, flac, gsm, h263, h264, hcom, ircam, la, lu, m4a, m4v, mat, mat4, mat5, maud, mov, mp2, mp3, mp4, mpeg, mpeg1, mpegts, mpg, nist, nul, ogg, paf, pcm, prc, pvf, raw, rawNVF, redir, rm, rtp, rtsp, sl6be, sl6le, s8, sb, sd2, sd2f, sdp, sf, sffd, shn, sl, smp, snd, sndt, sph, svx, sw, txw, u16be, u16le, u8, ub, ul, uw, vms, vob, voc, vorbis, vox, w64, wav, wma, wve, xi

Výstupní formáty:

aif, aiff, au, flac, gsm, m4a, mp3, ogg, raw, vox, wav, 8svx, ac3, aif, aifc, aiff, al, ape, asf, au, avr, cdr, cvs, dat, fap, ffm, flac, gsm, hcom, ircam, la, lu, m4a, mat, mat4, mat5, maud, mp3, mp4, mpeg, nist, ogg, paf, prc, pvf, raw, rm, sb, sf, shn, sl, smp, snd, sndt, sph, svx, sw, txw, ub, ul, uw, vms, vob, voc, vox, w64, wav, wve

## 16. Závěr

Diplomové práce byla zaměřena na tvorbu a zpracování signálových databází.

První část práce byla věnována obecné studii o vzniku databází, její historii a rozdělení na databázové modely. Dále stručně rozebírá, čím se vyznačuje multimediální databáze. Pojednává o dostupných databázích pro zpracování hudebních a řečových signálů. Bylo nalezeno několik softwarových prostředí k vytváření vlastních databázových struktur.

V další části byla věnována pozornost databázi SUSAS, která shromažďuje nahrávky, které byly pořízeny pod skutečným a simulovaným stresem. Nahrávky skutečného stresu byly pořízeny nahráváním řečových projevů pilotů při letovém provozu v krizových situacích, kdy pilot musel reagovat na neočekávané události. Tato databáze zkoumá rozdíly mezi těmito autentickými nahrávkami a obdobnými, které však simulovaně namluvili herci. K databázi SUSAS byla přeložena česká dokumentace a byly zpracovány její řečové soubory, které byly převedeny do standardního zvukového formátu wav.

Dále byla vytvořena banka hudebních a řečových signálů. Tato banka obsahuje mnoho kategorií a byla seříděna do logické struktury. Jednotlivé kategorie se vyznačují určitou podobností. Je zkoumána i jejich podobnost v grafickém vyjádření. Bylo pojednáno o různých hudebních formátech.

Hudební databáze obsahuje dvě kategorie. Melodie různými nástroji obsahuje tři známé melodie, hrané na různé nástroje v různých tóninách, reálnými i virtuálními nástroji, či různými technikami hraní. Další kategorie obsahuje stejné písně ve více žánrech. Zde jsou shromážděny nahrávky stejné písně, avšak zahrané vždy jiným hudebním umělcem, nebo při studiové a koncertní verzi.

Řečová databáze obsahuje řečové záznamy z databáze SUSAS a dále nahrávky z televizních pořadů, reality-show, sportovních přenosů, reportáží a dokumentů, kdy mluvčí je vystaven stresu a emocím.

Výsledkem diplomové práce je vytvoření vlastního databázového programu, který účinně třídí a zpracovává vlastní zvukové signály. V tomto programu si všechny záznamy můžeme přehrát, zobrazit u nich spektrogram a časový průběh až u třech záznamů současně z důvodu porovnání mezi sebou. Je možné uložit konkrétní záznam, či celou skupinu jako zvukový soubor. Databázi je možné volně rozšiřovat. Výsledný program byl zrealizován v české i anglické verzi.

## 17. Seznam literatury

- [1] Databázový svět  
<http://www.dbsvet.cz>
- [2] Wikipedia  
<http://cs.wikipedia.org>
- [3] Google  
<http://www.google.cz>
- [4] Encyklopedie Navajo  
<http://www.navajo.cz>
- [5] Webové stránky jednotlivých produktů  
<http://www.oracle.com>  
<http://weddlesoftware.com/AudioInfoDB.htm>  
<http://www.musicdatabasesoftware.com>  
<http://www.wakefieldsoft.com/audiolis>  
<http://www.filesland.com/companies/FNProgramvare/CATraxx.html>  
<http://www.brilliantdatabase.com>  
<http://www.websitedatabases.com>  
<http://www.lightdb.com>
- [6] Převod audio formátů  
[http://library.rice.edu/services/digital\\_media\\_center/online\\_guides/working-with-linguistic-resources/converting-sph-audio-files-to-wav](http://library.rice.edu/services/digital_media_center/online_guides/working-with-linguistic-resources/converting-sph-audio-files-to-wav)
- [7] Převod audio formátů  
<http://www.dekorte.com/projects/shareware/SoundConverter/>
- [8] Databáze Firebird  
<http://www.firebirdsql.org/>
- [9] GLETT, J. *Zpracování akustických signálů s multimediální prezentací*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2007. 42 s. Vedoucí bakalářské práce doc. Ing. Milan Sigmund, CSc.
- [10] Audio formáty  
[http://technet.idnes.cz/tec\\_technika.asp?r=digital&c=A000326\\_0012928\\_digital](http://technet.idnes.cz/tec_technika.asp?r=digital&c=A000326_0012928_digital)
- [11] Audio formáty  
<http://www.pixel.cz/pdf/129-36.pdf>

## 18. Zdroje zvukových signálů a použité hudební programy

Zvuky „Melodie různými nástroji“ byly použity z práce:

GLETT, J. *Zpracování akustických signálů s multimediální prezentací*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2007. 42 s. Vedoucí bakalářské práce doc. Ing. Milan Sigmund, CSc.

Soubory v kapitole „Stejná píseň ve více žánrech“:

Studiové nahrávky byly převzaty z originálních alb umělců, verze s označením „youtube.com“ byli pořízeny ze stejnojmenného serveru a „živé nahrávky“ jsem pořídil při živém vystoupení umělců na koncertech, které jsem zvučil.

K nahrávání, generování a dalšímu zpracování jsem použil následující hudební programy:

SIA SmaartLive 5

- vytváření spektrogramů

(bylo použito nastavení pro delší signály: Scale: 1/24, Avg: 2, Mag Th: Off, Weight: None, SR: 44100, FFT: 16k, TC: 372ms, FR: 2,7Hz, Spectro range: -20 až -70 dB

pro krátké signály (do cca 1s.): Scale: 1/24, Avg: 2, Mag Th: Off, Weight: None, SR: 44100, FFT: 32k, TC: 743ms, FR: 1,3Hz, Spectro range: -20 až -70 dB)

<http://www.eaw.com/products/software/EAWSmaart/index.html>

Cubase System 4

- generování a nahrávání, notové partitury

[http://www.steinberg.net/en/products/musicproduction/cubase4\\_product.html](http://www.steinberg.net/en/products/musicproduction/cubase4_product.html)

Steinberg Wavelab 6

- normalizace

[http://www.steinberg.net/en/products/audioediting\\_product/audioediting\\_wavelab6.html](http://www.steinberg.net/en/products/audioediting_product/audioediting_wavelab6.html)

Sony Sound Forge 9.0

- nahrávání z youtube.com a vytváření časových průběhů

<http://www.sonycreativesoftware.com/soundforge>

virtuální midi nástroje firmy Steinberg

- zdrojové midi banky pro virtuální nástroje

<http://www.steinberg.net/en/products/vst instruments.html>

Sph convert v2.1

- převod formátů sph -> wav

[http://library.rice.edu/services/digital\\_media\\_center/online\\_guides/working-with-linguistic-resources/converting-sph-audio-files-to-wav](http://library.rice.edu/services/digital_media_center/online_guides/working-with-linguistic-resources/converting-sph-audio-files-to-wav)

CdeX 1.70

- převod formátů mp3 -> wav

<http://cdexos.sourceforge.net/>